

# Submodular Cost Allocation Problem and Applications\*

Chandra Chekuri<sup>†</sup>

Alina Ene<sup>‡</sup>

January 20, 2013

## Abstract

We study the Minimum Submodular-Cost Allocation problem (MSCA). In this problem we are given a finite ground set  $V$  and  $k$  non-negative submodular set functions  $f_1, \dots, f_k$  on  $V$ . The objective is to partition  $V$  into  $k$  (possibly empty) sets  $A_1, \dots, A_k$  such that the sum  $\sum_{i=1}^k f_i(A_i)$  is minimized. Several well-studied problems such as the non-metric facility location problem, multiway-cut in graphs and hypergraphs, and uniform metric labeling and its generalizations can be shown to be special cases of MSCA. In this paper we consider a convex-programming relaxation obtained via the Lovász-extension for submodular functions. This allows us to understand several previous relaxations and rounding procedures in a unified fashion and also develop new formulations and approximation algorithms for several problems. In particular, we give a  $(1.5 - 1/k)$ -approximation for the hypergraph multiway partition problem. We also give a  $\min\{2(1 - 1/k), H_\Delta\}$ -approximation for the hypergraph multiway cut problem when  $\Delta$  is the maximum hyperedge size. Both problems generalize the multiway cut problem in graphs and the hypergraph cut problem is approximation equivalent to the node-weighted multiway cut problem in graphs.

---

\*An extended abstract of this paper will appear in *Proc. of ICALP*, July 2011.

<sup>†</sup>Dept. of Computer Science, University of Illinois, Urbana, IL 61801. Supported in part by NSF grants CCF-0728782 and CCF-1016684. [chekuri@cs.illinois.edu](mailto:chekuri@cs.illinois.edu)

<sup>‡</sup>Dept. of Computer Science, University of Illinois, Urbana, IL 61801. Supported in part by NSF grants CCF-0728782 and CCF-1016684. [ene1@illinois.edu](mailto:ene1@illinois.edu)

# 1 Introduction

We consider the following allocation problem with submodular costs.

**Minimum Submodular-Cost Allocation (MSCA).** Let  $V$  be a finite ground set and let  $f_1, \dots, f_k$  be  $k$  non-negative submodular set functions on  $V$ . That is, for  $1 \leq i \leq k$ ,  $f_i : 2^V \rightarrow \mathbb{R}_+$  and  $f_i(A) + f_i(B) \geq f_i(A \cup B) + f_i(A \cap B)$  for all  $A, B \subseteq V$ . In the MSCA problem the goal is to partition the ground set  $V$  into  $k$  (possibly empty) sets  $A_1, \dots, A_k$  such that the sum  $\sum_{i=1}^k f_i(A_i)$  is minimized.

We observe that the problem is interesting only if the  $f_i$ 's are different for otherwise allocating all of  $V$  to  $f_1$  is trivially an optimal solution. We assume that the functions  $f_i$  are given as value oracles, although in specific applications they may be available as explicit poly-time computable functions of some auxiliary input. The special case of this problem in which all of the functions are monotone ( $f(A) \leq f(B)$  if  $A \subseteq B$ ) has been previously considered by Svitkina and Tardos [22]. In this paper, we consider the problem with both monotone and non-monotone functions. We show that several well-studied problems such as non-metric facility location, multiway cut problems in graphs and hypergraphs, uniform metric labeling and its generalization to hub location among others can be cast as special cases of MSCA. In particular, we investigate the integrality gap of a simple and natural convex-programming relaxation for MSCA that is obtained via the use of the Lovász extension of a submodular function.

**Lovász extension and a convex program for MSCA:** Let  $V$  be a finite ground set of cardinality  $n$ . Each real-valued set function on  $V$  corresponds to a function  $f : \{0, 1\}^n \rightarrow \mathbb{R}$  on the vertices of the  $n$ -dimensional hypercube. The Lovász extension of  $f$  to the continuous domain  $[0, 1]^n$  denoted by  $\hat{f}$  is defined as<sup>1</sup>

$$\hat{f}(\mathbf{x}) = \mathbb{E}_{\theta \in [0, 1]} [f(\mathbf{x}^\theta)] = \int_0^1 f(\mathbf{x}^\theta) d\theta$$

where  $\mathbf{x}^\theta \in \{0, 1\}^n$  for a given vector  $\mathbf{x} \in [0, 1]^n$  is defined as:  $x_i^\theta = 1$  if  $x_i \geq \theta$  and 0 otherwise.

Lovász showed that  $\hat{f}$  is convex if and only if  $f$  is a submodular set function [17]. Moreover, it is easy to see that, given  $\mathbf{x}$ , the value  $\hat{f}(\mathbf{x})$  can be computed in polynomial time by using a value oracle for  $f$ . Via this extension, we obtain a straightforward relaxation for MSCA with a convex objective function and linear constraints. Let  $v_1, \dots, v_n$  denote the elements of  $V$ . The relaxation has variables  $x(v, i)$  for  $v \in V$  and  $1 \leq i \leq k$  with the interpretation that  $x(v, i)$  is 1 if  $v$  is assigned to  $A_i$  and 0 otherwise. Let  $\mathbf{x}_i = (x(v_1, i), \dots, x(v_n, i))$ . The relaxation is given below.

<b>LE-REL</b>	
min	$\sum_{i=1}^k \hat{f}_i(\mathbf{x}_i)$
	$\sum_{i=1}^k x(v, i) = 1 \quad \forall v$
	$x(v, i) \geq 0 \quad \forall v, i$

Throughout, we use  $\text{OPT}$  and  $\text{OPT}_{\text{FRAC}}$  to denote the value of an optimal integral and an optimal fractional solution to LE-REL (respectively).

We remark that LE-REL can be solved in time that is polynomial in  $n$  and  $\log(\max_{S \subseteq V} f(S))$  via the ellipsoid method; we give some of the details in Appendix A. Moreover, for some problems of interest the

---

<sup>1</sup>The definition is not the standard one but is equivalent to it; see [24] or Appendix A. This definition is convenient to us in describing and understanding rounding procedures.

above convex program can be rewritten into an equivalent linear program. We now describe several problems that can be cast as special cases of MSCA, and also how some previously considered linear-programming relaxations can be seen as being equivalent to the convex program above.

## 1.1 Problems related to MSCA

**Monotone MSCA (MONOTONE-MSCA) and Facility Location:** In facility location, we have a set of facilities  $\mathcal{F}$  and a set of clients or demands  $\mathcal{D}$ . There is a non-negative cost  $c_{ij}$  to connect facility  $i$  to client  $j$  (we do not necessarily assume that these costs form a metric). Opening facility  $i \in \mathcal{F}$  costs  $f_i$ . The goal is to open a subset of the facilities and assign each client to an open facility so as to minimize the sum of the facility opening cost and the connection costs. Svitkina and Tardos [22] considered the setting where the cost of opening a facility  $i$  is a monotone submodular function  $g_i$  of the clients assigned to it, and gave an  $(1 + \ln |\mathcal{D}|)$ -approximation, and matching hardness via a reduction from set cover. We note that this problem is equivalent to MSCA when all the  $f_i$  are monotone submodular functions, which we refer to as MONOTONE-MSCA. In [22] a greedy algorithm via submodular function minimization is used to derive the approximation. Here we prove that the integrality gap of LE-REL is  $(1 + \ln |\mathcal{D}|)$ , and describe how certain rounding algorithms achieve this bound. These algorithms are useful when considering functions that are not necessarily monotone.

**Submodular Multiway Partition (SUB-MP):** We define an abstract problem and then specialize to known problems. Let  $f : 2^V \rightarrow \mathbb{R}_+$  be a submodular set function over  $V$  and let  $S = \{s_1, s_2, \dots, s_k\}$  be  $k$  terminals in  $V$ . The submodular multiway partition problem is to find a partition of  $V$  into  $A_1, \dots, A_k$  such that  $s_i \in A_i$  and  $\sum_{i=1}^k f(A_i)$  is minimized. This has been previously considered by Zhao, Nagamochi and Ibaraki [27]. This can be seen as a special case of MSCA as follows. Define the ground set to be  $V' = V \setminus S$  and, for  $1 \leq i \leq k$ ,  $f_i : 2^{V'} \rightarrow \mathbb{R}_+$  is the function defined as  $f_i(S) = f(S \cup \{s_i\})$ . If in addition  $f$  is symmetric ( $f(A) = f(V - A)$  for all  $A$ ) we call this the symmetric SUB-MP problem (SYM-SUB-MP). Note that although the problem is based on a single function  $f$ ,  $k$  different submodular functions (induced by the terminals) are needed to reduce it to MSCA. We now discuss some important special cases of this problem.

**Multiway Cut in Graphs (GRAPH-MC):** The input is an edge-weighted undirected graph  $G = (V, E)$  and  $k$  terminal vertices  $S = \{s_1, \dots, s_k\}$ ; the goal is to remove a minimum-weight set of edges to disconnect the terminals. This can be seen as a special case of the symmetric submodular multiway partition problem by simply choosing  $f$  to be the cut-capacity function of  $G$  scaled down by a factor of 2. That is,  $f(A) = \frac{1}{2} \sum_{e \in \delta(A)} w(e)$  where  $w(e)$  is the weight of edge  $e$ . We observe that LE-REL for this problem is equivalent to the well-known geometric LP relaxation of Calinescu, Karloff and Rabani [2], which led to significant improvements ( $1.5 - 1/k$  in [2] and 1.3438 in [14]) over the  $2(1 - 1/k)$ -approximation obtained via the isolating-cut heuristic [4].

**Multiway Cut and Partition in Hyper-Graphs:** Given an edge-weighted hypergraph  $\mathcal{G} = (V, \mathcal{E})$  and terminal set  $S \subset V$ , the HYPERGRAPH MULTIWAY CUT problem (HYPERGRAPH-MC) (see [18, 26, 8]) asks for the minimum weight subset of hyperedges whose removal disconnects the terminals. This can be seen as a special case of SUB-MP [18]; this reduction requires some care and the underlying submodular function is *asymmetric*. A related problem is the HYPERGRAPH MULTIWAY PARTITION problem (HYPERGRAPH-MP) introduced by Lawler [16] where the cost for hyperedge  $e$  is proportional to the number of non-trivial pieces it is partitioned into. This can be seen as a special case of the SYM-SUB-MP with  $f$  being the hypergraph cut capacity function. We note that GRAPH-MC is a special case of both HYPERGRAPH-MC and HYPERGRAPH-MP.

**Node-weighted Multiway Cut in Graphs (NODE-WT-MC):** In this problem [9] the graph has weights on nodes instead of edges and the goal is to find a minimum weight subset of nodes whose removal disconnects a given set of terminals. It is not difficult to show that HYPERGRAPH-MC and NODE-WT-MC are approximation equivalent [18].

Zhao *et al.* [27] consider generalizations of the above problems where some set of terminals  $S \subseteq V$  and  $k$  are specified and the goal is to partition  $V$  into  $k$  sets such that each set contains at least one terminal and the

total cost of the partition is minimized. We do not discuss these further since they are not directly related to MSCA, although one can reduce them to MSCA if  $k$  is a fixed constant.

**Uniform Metric Labeling and SUBMODULAR COST LABELING (SUB-LABEL):** The metric labeling problem was introduced by Kleinberg and Tardos [15] as a general classification problem. We are given an undirected edge-weighted graph  $G = (V, E)$  and  $k$  labels and the goal is to assign a label to each vertex to minimize the labeling cost and the edge-cut cost. Assigning label  $i$  to  $v$  incurs a cost  $c_i(v)$  and if an edge  $uv$  of weight  $w(uv)$  has  $u$  labeled with  $i$  and  $v$  labeled with  $j$  then the edge-cut cost incurred is  $w(uv) \cdot d(ij)$ . The uniform metric labeling problem is obtained when  $d(ij) = 1$  for all  $i \neq j$ . We consider the following generalization that we call the SUBMODULAR COST LABELING (SUB-LABEL) problem which is a special case of MSCA. The  $k$  labels correspond to the  $k$  functions  $f_1, \dots, f_k$ . We define  $f_i$  as the sum of two functions, a monotone function  $g_i$  that models the label assignment cost, and a non-monotone function  $h$  that models the cut-cost. The goal then is to partition  $V$  into  $A_1, \dots, A_k$  to minimize  $\sum_{i=1}^k (g_i(A_i) + h(A_i))$ . Note that uniform metric labeling is the special case when  $g_i$  are modular and  $h$  is the graph cut function, which is symmetric. We are motivated to consider this generalization by problems that have been considered previously, such as metric labeling on hypergraphs, hub location problem [10], and the extension of metric labeling to handle label opening costs [5].

## 1.2 Overview of Results and Techniques

In this paper we examine the complexity of MSCA primarily through the “integrality gap” of the convex relaxation LE-REL which can be optimized in polynomial time. All the problems we consider are NP-hard and our focus is on polynomial time approximation algorithms.

A significant portion of our contribution is to highlight the naturalness of MSCA and the Lovász-extension based relaxation LE-REL by showing connections to previously studied problems, linear programming relaxations, and rounding strategies. Viewing these problems in the more abstract setting of submodularity gives insights into prior algorithms. In the process, we obtain new and interesting results. Although one would like to obtain a single unifying algorithm that achieves a good approximation for MSCA, it turns out that LE-REL has a large integrality gap and we believe that MSCA is hard to approximate to a polynomial factor. However, it is fruitful to examine special cases of MSCA that admit good approximations via LE-REL. We describe several applications below by summarizing our results; all of them are based on LE-REL.

- The integrality gap of LE-REL for MONOTONE-MSCA is  $\Theta(\log n)$ .
- There is a  $(1.5 - 1/k)$ -approximation for HYPERGRAPH-MP.
- There is a  $\min\{2(1 - 1/k), H_\Delta\}$ -approximation for HYPERGRAPH-MC, where  $\Delta$  is the maximum hyperedge size and  $H_i$  is the  $i$ -th harmonic number. For  $\Delta = 2$  this gives a 1.5-approximation and for  $\Delta = 3$  this gives a 1.833-approximation.
- LE-REL for HYPERGRAPH-MC gives a new mathematical programming relaxation for NODE-WT-MC and a new 2-approximation. Moreover, if all non-terminal nodes have degree at most 3 we obtain a 1.833-approximation improving upon the  $2(1 - 1/k)$  known via the distance-based relaxation [9].
- The integrality gap of LE-REL for SYM-SUB-MP is at most  $2 - 2/k$ ; this gives an alternative approximation to previous combinatorial algorithms [19, 27]. We raise the question as to whether the integrality gap is at most 1.5.
- There is an  $O(\log n)$  for SUB-LABEL when the cut function is symmetric. We derive results for other special cases of SUB-LABEL.

**Rounding the convex relaxation:** Recall that the objective function in LE-REL is  $\sum_{i=1}^k \hat{f}_i(\mathbf{x}_i)$ , where  $\hat{f}_i(\mathbf{x}_i) = \mathbb{E}_{\theta \in [0,1]} [f(\mathbf{x}_i^\theta)]$ . How do we round while preserving the objective function? If we focus on a specific  $i$ , the objective function suggests that we pick  $\theta$  randomly from  $[0, 1]$  and assign the elements in  $\mathbf{x}_i^\theta$  to  $i$ ; we call this  $\theta$ -rounding. However, there are two issues to contend with. First, if we independently round for each  $i$  then the same element may be assigned multiple times. Second, we need to ensure that all elements are assigned, which is not guaranteed by the  $\theta$ -rounding. We remark that there is an integrality gap example for hypergraph metric labeling that shows that there is no effective rounding strategy that works in general.

Our approach is to understand the rounding process by considering various special cases of interest. In particular, we consider monotone functions, symmetric functions, the hypergraph separation cost function (which is asymmetric), and combinations of such functions. Monotonicity helps in that if elements are assigned to a label  $i$ , they can be removed without increasing the fractional cost. Although one can use different strategies to obtain an  $O(\log n)$ -approximation and integrality gap, a useful strategy here is the rounding of Kleinberg and Tardos [15] that they introduced for metric labeling. This has the additional property of ensuring that an element  $u$  is assigned to  $i$  with probability exactly  $x(u, i)$ . We then consider the rounding process for SUB-MP, in particular the symmetric case SYM-SUB-MP. Here, we crucially take advantage of the fact that there is a single underlying function  $f$ , and moreover the fact that it is symmetric. We consider the CKR-ROUNDING strategy from [2] and show its effectiveness for hypergraphs by abstracting away some of the properties specific to graphs that were previously exploited in the analysis. In the process, we also observe that a variant is equally effective for graphs but is more insightful for SYM-SUB-MP.

Finally, SUB-LABEL combines a monotone function and a non-monotone function. Here, we resort to KT-ROUNDING since it is a reasonable strategy to approximately preserve the cost of the monotone component. For the uniform metric labeling problem, [15] showed that KT-ROUNDING approximately (to within a factor of 2) preserves the fractional connection cost in the case of graphs. We show bounds for hypergraph cut functions in an analogous fashion. Our insights enable us to develop a variant of the rounding that gives an  $O(\log n)$ -approximation for SUB-LABEL when the cut function is an arbitrary symmetric submodular function.

**Other Related Work:** There has been much recent interest in optimizing with submodular set functions. In particular, maximization problems have been examined via combinatorial techniques as well as the multilinear relaxation [1]. The submodular welfare problem [23] is similar in spirit to MSCA except that one is interested in maximizing the value of an allocation rather than minimizing the cost. Minimization problems with submodular costs have also received substantial attention [20, 12, 13, 11] with several negative results for basic problems as well as positive approximation results for problems such as the submodular cost vertex cover problem [13, 11]. Lovász-extension based convex programs have been effectively used for these problems. Various submodular cut and partition problems and their special cases such as the hypergraph cut and partition have been studied recently [27, 26, 18, 8]; however, these papers have typically focussed on greedy and divide-and-conquer based approaches while we use LE-REL.

**Recent Results for SYM-SUB-MP and SUB-MP:** Very recently, building on the work in this paper and a non-trivial new technical theorem, we showed [3] that the integrality gap of SUBMP-REL is at most  $1.5 - 1/k$  for SYM-SUB-MP and at most 2 for SUB-MP.

## 2 Monotone MSCA

In this section we consider MONOTONE-MSCA where  $f_1, \dots, f_k$  are monotone submodular functions. We will assume for simplicity that  $f_i(\emptyset) = 0$  for all  $i$ . Svitkina and Tardos [22] considered this problem in the context of facility location and gave a  $(1 + \ln n)$ -approximation and matching hardness via an approximation preserving reduction from set cover. Let  $\alpha = \min_{S \subseteq V, 1 \leq i \leq k} f_i(S)/|S|$ . The main observation in [22] is that  $\alpha \leq \text{OPT}/n$ , and moreover a pair  $(S, i)$  such that  $f_i(S)/|S| = \alpha$  can be computed in polynomial time via

submodular function minimization. One can then iterate using a greedy scheme, by using the monotonicity of the functions, to obtain a  $(1 + \ln n)$ -approximation. Using a similar argument, we can prove the following theorem.

**Theorem 2.1.** *The integrality gap of LE-REL for MONOTONE-MSCA is at most  $(1 + \ln n)$ . In particular,  $\alpha \leq \text{OPT}_{\text{FRAC}}/n$ .*

**MONOTONE-MSCA-GREEDY:**

```

let  $x$  be a solution to LE-REL
 $A_i \leftarrow \emptyset$  for all  $i$  ( $1 \leq i \leq k$ ) «the set of vertices that will be assigned to  $i$ »
 $U \leftarrow V$  «the set of unassigned vertices»
while  $U \neq \emptyset$ 
  let  $\tilde{x}$  be the restriction of  $x$  to  $U$ 
  for each  $\theta$ , let  $A(i, \theta) = \{v \mid v \in U, \tilde{x}(v, i) \geq \theta\}$ 
  let  $0 = \theta_{i,1} < \theta_{i,2} < \dots < \theta_{i,\ell_i} = 1$  be the distinct entries of  $\tilde{\mathbf{x}}_i$ 
  let  $(i', j')$  be the pair of indices in the set  $\{(i, j) \mid 1 \leq i \leq k, 1 \leq j < \ell_i\}$ 
    that minimizes the ratio  $f_i(A(i, \theta_{i,j})) / |A(i, \theta_{i,j})|$ 
   $A_{i'} \leftarrow A_{i'} \cup A(i', \theta_{i',j'})$ 
   $U \leftarrow U - A(i', \theta_{i',j'})$ 

```

**Theorem 2.2.** MONOTONE-MSCA-GREEDY achieves an  $H_n$ -approximation for MONOTONE-MSCA.

Before we prove Theorem 2.2, we introduce some notation. Consider iteration of MONOTONE-MSCA-GREEDY. Consider an iteration of the while loop of MONOTONE-MSCA-GREEDY. Let  $U$  be the set of elements that are unassigned at the beginning of the iteration, and let  $\tilde{x}$  denote the restriction of  $x$  to  $U$ ; more precisely,  $\tilde{x}(v, i) = x(v, i)$  for all terminals  $i$  and all vertices  $v \in U$ . For any  $\theta$ , Let  $A(i, \theta) = \{v \mid v \in U, \tilde{x}(v, i) \geq \theta\}$ . Let  $0 = \theta_{i,1} < \theta_{i,2} < \dots < \theta_{i,\ell_i} = 1$  be the distinct entries of  $\tilde{\mathbf{x}}_i$ . Let  $\text{OPT}_{\text{FRAC}} = \sum_{i=1}^k f_i(\mathbf{x}_i)$ . Theorem 2.2 follows from the following lemma.

**Lemma 2.3.**

$$\min_{1 \leq i \leq k} \min_{0 \leq \theta < \ell_i} \frac{f_i(A(i, \theta_{i,j}))}{|A(i, \theta_{i,j})|} \leq \frac{\text{OPT}_{\text{FRAC}}}{|U|}.$$

In order to prove Lemma 2.3, we will show that, if we choose a terminal  $i \in \{1, 2, \dots, k\}$  and a threshold  $\theta \in [0, 1]$  uniformly at random, the ratio  $\mathbb{E}[f_i(A(i, \theta))] / \mathbb{E}[|A(i, \theta)|]$  is at most  $\text{OPT}_{\text{FRAC}}/|U|$ . The following propositions give a bound on the two expectations  $\mathbb{E}[f_i(A(i, \theta))]$  and  $\mathbb{E}[|A(i, \theta)|]$ ; their proofs are relatively straightforward and they have been moved to Appendix B.

**Proposition 2.4.**

$$\mathbb{E}_{i, \theta}[f_i(A(i, \theta))] \leq \frac{1}{k} \text{OPT}_{\text{FRAC}}.$$

**Proposition 2.5.**

$$\mathbb{E}_{i, \theta}[|A(i, \theta)|] = \frac{1}{k} |U|.$$

**Proof of Lemma 2.3:** Let  $i$  be a terminal selected uniformly at random. Let  $\theta$  be a threshold selected uniformly at random from the interval  $[0, 1]$ . It follows from Proposition 2.4 and Proposition 2.5 that the ratio  $\mathbb{E}[f_i(A(i, \theta))] / \mathbb{E}[|A(i, \theta)|]$  is at most  $\text{OPT}_{\text{FRAC}}/|U|$ . By linearity of expectation,

$$\mathbb{E}_{i, \theta} \left[ f_i(A(i, \theta)) - \frac{\text{OPT}_{\text{FRAC}}}{|U|} \cdot |A(i, \theta)| \right] \leq 0$$

and therefore there exists a terminal  $i'$  and a threshold  $\theta'$  for which the ratio  $f_{i'}(A(i', \theta'))/|A(i', \theta')|$  is at most  $\text{OPT}_{\text{FRAC}}/|U|$ . Let  $j'$  be the smallest index  $j$  that satisfies  $0 \leq j < \ell_{i'}$  and  $\theta_{i',j} \geq \theta'$ . Since  $A(i', \theta_{i',j'}) = A(i', \theta')$ ,  $(i', j')$  is the desired pair.  $\square$

In the remainder of this section, we consider a different algorithm that achieves an  $O(\log n)$ -approximation for MONOTONE-MSCA. We will use this algorithm as a building block for submodular cost labeling algorithms (see Section 4). The algorithm KT-ROUNDING is derived from the work of Kleinberg and Tardos on metric labeling [15].

#### KT-ROUNDING

```

let  $x$  be a solution to LE-REL
 $S \leftarrow \emptyset$        $\langle\langle$ set of all assigned vertices $\rangle\rangle$ 
 $\langle\langle$ set of vertices that are eventually assigned to  $i$  $\rangle\rangle$ 
 $A_i \leftarrow \emptyset$  for all  $i$  ( $1 \leq i \leq k$ )
while  $S \neq V$ 
    pick  $i \in \{1, 2, \dots, k\}$  uniformly at random
    pick  $\theta \in [0, 1]$  uniformly at random
     $A_i \leftarrow A_i \cup (\{v \mid x(v, i) \geq \theta\} - S)$ 
     $S \leftarrow S \cup A_i$ 
return  $(A_1, \dots, A_k)$ 

```

We prove the following theorem by building on some useful properties that are shown in [15]; one of these is that the probability that  $v$  gets assigned to  $i$  in the rounding is precisely  $x(v, i)$ . In particular, this yields an optimal algorithm for modular functions.

**Theorem 2.6.** *KT-ROUNDING achieves a randomized  $O(\ln n)$ -approximation for MONOTONE MSCA.*

**Proof Sketch:** It is shown in [15], and not difficult to see, that the rounding terminates in  $O(k \log n)$  iterations of the while loop with high probability. In each iteration the algorithm does a  $\theta$ -rounding on an index chosen uniformly at random. Let  $i$  be the random index and  $A(i, \theta) = \{v \mid x(v, i) \geq \theta\}$ . Then it is easy to see that  $\mathbb{E}[f_i(A(i, \theta))] = \sum_{i=1}^k \frac{1}{k} \hat{f}_i(\mathbf{x}_i) = \frac{1}{k} \text{OPT}_{\text{FRAC}}$ . Since the functions are monotone, we have that  $\mathbb{E}[f(A(i, \theta) - S)] \leq \frac{1}{k} \text{OPT}_{\text{FRAC}}$ . Since the algorithm terminates in  $O(k \log n)$  iterations, by linearity of expectation and the sub-additivity of the functions (since the functions are submodular and  $f(\emptyset) = 0$ ), the total expected cost is  $O(\log n) \text{OPT}_{\text{FRAC}}$ .  $\square$

### 3 Submodular Multiway Partition

We consider MSCA when the  $f_i$  can be non-monotone. We can show that the integrality gap of LE-REL even for a special case of labeling on hypergraphs can be  $\Omega(n)$ , and we suspect that the problem is hard to approximate to a polynomial factor in  $n$ . We therefore focus on SUBMODULAR MULTIWAY PARTITION (SUB-MP) and SUBMODULAR COST LABELING (SUB-LABEL); these are broad special cases which capture several problems that have been considered previously.

The reduction of SUB-MP to MSCA requires one to work with the non-terminals  $V'$  as the ground set. It is however more convenient to work with the terminals and non-terminals. In particular, we work with the relaxation below. Recall that  $\mathbf{x}_i = (x(v_1, i), \dots, x(v_n, i))$ .

<b>SUBMP-REL</b>			
min	$\sum_{i=1}^k \hat{f}(\mathbf{x}_i)$		
	$\sum_{i=1}^k x(v, i)$	$= 1$	$\forall v$
	$x(s_i, i)$	$= 1$	$\forall i$
	$x(v, i)$	$\geq 0$	$\forall v, i$

As before, a starting point for rounding the relaxation is the basic  $\theta$ -rounding that preserves the objective function. Suppose we do  $\theta$ -rounding for each  $i$  to obtain sets  $A(1, \theta), \dots, A(k, \theta)$  where each  $A(i, \theta) \subseteq V$ . Here we could use independent random  $\theta$  values for each  $i$  or the same  $\theta$ . Note that the constraints ensure that  $s_i \in A(j, \theta)$  iff  $i = j$ . However, the sets  $A(1, \theta), \dots, A(k, \theta)$  may intersect and also may not cover the entire set  $V$ , in which case we have to allocate the remaining elements in some fashion. First we show how to take advantage of the case when  $f$  is symmetric and then discuss how to obtain results for hypergraph problems that are special cases of SUB-MP.

**A  $2(1 - 1/k)$ -approximation for SYM-SUB-MP:** A  $2(1 - 1/k)$ -approximation for SYM-SUB-MP is known via greedy combinatorial algorithms [19, 27]. However, no mathematical programming formulation for the problem has been previously considered. Here we show that, on instances of SYM-SUB-MP, the integrality gap of LE-REL is  $2(1 - 1/k)$  by using an uncrossing property of symmetric functions.

The following lemma is standard and it has been used in previous work [21].

**Lemma 3.1.** *Let  $f$  be a symmetric submodular set function over  $V$  and let  $A_1, \dots, A_k$  be subsets of  $V$ . Then there exist sets  $A'_1, \dots, A'_k$  such that (i)  $A'_i \subseteq A_i$  for  $1 \leq i \leq k$ , (ii)  $A'_1, \dots, A'_k$  are mutually disjoint (iii)  $\cup_i A'_i = \cup_i A_i$  and (iv)  $\sum_i f(A'_i) \leq \sum_i f(A_i)$ . Moreover, given the  $A_i$ 's a collection of sets  $A'_i$  satisfying the above properties can be found in polynomial time via a value oracle for  $f$ .*

**Proof:** Since  $f$  is symmetric, it satisfies posi-modularity; that is,  $f(X) + f(Y) \geq f(X - Y) + f(Y - X)$ . From this we see that either  $f(X) + f(Y - X)$  or  $f(Y) + f(X - Y)$  is no larger than  $f(X) + f(Y)$ . This allows us to uncross  $A_1, \dots, A_k$  as follows. If the  $A_i$ 's are mutually disjoint then we can set  $A'_i = A_i$  for each  $i$  and they satisfy the desired properties. Otherwise, there exist distinct  $i$  and  $j$  such that  $A_i \cap A_j \neq \emptyset$ . We can replace  $A_i$  and  $A_j$  with  $A_i$  and  $A_j - A_i$  if  $f(A_i) + f(A_j - A_i) \leq f(A_i) + f(A_j)$ ; otherwise, we replace them by  $A_i - A_j$  and  $A_j$ . We repeat this process to get the desired sets.  $\square$

**Theorem 3.2.** *The integrality gap of LE-REL for SYM-SUB-MP is  $\leq 2(1 - 1/k)$ .*

**Proof:** Let  $\mathbf{x}$  be an optimal solution to LE-REL for a given instance of SYM-SUB-MP. Let  $A(1, \theta), \dots, A(k, \theta)$  be sets obtained by applying  $\theta$ -rounding to each  $i$ . By the property of  $\theta$ -rounding, we observe that  $\sum_i \mathbb{E}[f(A(i, \theta))] = \sum_i \hat{f}(\mathbf{x}_i) = \text{OPT}_{\text{FRAC}}$ . Note that  $s_i$  belongs only to  $A(i, \theta)$ . We now apply Lemma 3.1 to  $A(1, \theta), \dots, A(k, \theta)$  to obtain  $A'_1, \dots, A'_k$ . We have  $\sum_i f(A'_i) \leq \sum_i f(A(i, \theta))$ . Let  $V' = V - \cup_i A'_i$ . By symmetry of  $f$ ,  $f(V') = f(\cup_i A'_i)$  and, since  $f$  is sub-additive,  $f(V') = f(\cup_i A'_i) \leq \sum_i f(A'_i) \leq \sum_i f(A(i, \theta))$ . We can allocate  $V'$  to any index  $i$  and the total cost of the allocation is upper bounded by  $f(V') + \sum_i f(A'_i) \leq 2 \sum_i f(A(i, \theta))$ . Thus the expected cost of the allocation is at most  $2\text{OPT}_{\text{FRAC}}$ . The allocation is feasible since  $s_i$  belongs only to  $A(i, \theta)$  and hence to  $A'_i$ . One can refine this argument slightly to obtain a  $2(1 - 1/k)$  bound; we give the details in Appendix C.  $\square$

In a previous version of this paper, we raised the following question.

**Question.** Is the integrality gap of LE-REL for SYM-SUB-MP at most 1.5?

As we already noted, we have shown in subsequent work [3] that the integrality gap is at most  $1.5 - 1/k$ .



**Rounding for HYPERGRAPH-MC and HYPERGRAPH-MP:** Calinescu *et al.* [2] gave a new geometric relaxation for GRAPH-MC, and a rounding procedure that gave a  $(1.5 - 1/k)$ -approximation; the integrality gap was subsequently improved to a bound of  $1.3438 - \varepsilon_k$  in [14], while the best known lower bound is  $8/(7 + 1/k - 1)$  [7]. Calinescu *et al.* [2] derived their relaxation as a way to improve the integrality gap of  $2(1 - 1/k)$  for a natural distance based linear programming relaxation; in fact, it often goes unnoticed that [2] shows the equivalence of their geometric relaxation to that of another relaxation obtained by adding valid strengthening constraints to the distance based relaxation. Interestingly, when we specialize MSCA to GRAPH-MC, LE-REL becomes the geometric relaxation of [2]! The rounding procedure in [2] can be naturally extended to rounding LE-REL for SUB-MP and we describe it below.

**CKR-ROUNDING**

```

let  $x$  be a solution to SUBMP-REL
pick a random permutation  $\pi$  of  $\{1, 2, \dots, k\}$ 
pick  $\theta \in [0, 1)$  uniformly at random
 $S \leftarrow \emptyset$     «set of all assigned vertices»
for  $i = 1$  to  $k - 1$ 
     $A_{\pi(i)} \leftarrow (\{v \mid x(v, \pi(i)) \geq \theta\} - S)$ 
     $S \leftarrow S \cup A_{\pi(i)}$ 
 $A_{\pi(k)} \leftarrow V - S$ 
return  $(A_1, \dots, A_k)$ 

```

CKR-ROUNDING uses the same  $\theta$  for all  $i$  and a random permutation, both of which are crucially used in the 1.5-approximation analysis for GRAPH-MC. In this paper we investigate CKR-ROUNDING and other roundings for HYPERGRAPH-MC and HYPERGRAPH-MP.

Although HYPERGRAPH-MC and HYPERGRAPH-MP appear similar, their objective functions are different. The objective of HYPERGRAPH-MC is to remove a minimum weight subset of hyperedges such that the terminals are separated, whereas the objective of HYPERGRAPH-MP is to minimize  $\sum_e w(e)p(e)$ , where  $p(e)$  is the number of non-trivial parts that  $e$  is partitioned into (a part is non-trivial if some vertex of  $e$  is in that part but not all of  $e$ ). For graphs we have that either  $p(e) = 0$  or  $p(e) = 2$ , and therefore the two problems HYPERGRAPH-MC and HYPERGRAPH-MP are equivalent; this is the reason why one can view GRAPH-MC as a partition problem as well. However, when the hyperedges can have size larger than 2, the objective function values are not related to each other (it is easy to see that the HYPERGRAPH-MP objective is always larger).

HYPERGRAPH-MP and HYPERGRAPH-MC have been studied for their theoretical interest and their applications. It is easy to see from its definition that HYPERGRAPH-MP is a special case of SYM-SUB-MP. It has been observed by a simple yet nice reduction [18] that HYPERGRAPH-MC is a special case of SUB-MP. In addition, it has been observed that HYPERGRAPH-MC is approximation-equivalent to the *node-weighted* multiway cut problem in graphs (NODE-WT-MC) [9].

We show that CKR-ROUNDING gives a  $(1.5 - 1/k)$ -approximation to HYPERGRAPH-MP and a tight  $H_\Delta$ -approximation for HYPERGRAPH-MC with maximum hyperedge size  $\Delta$ . Note that when  $\Delta = 2$ ,  $H_\Delta = 1.5$  and when  $\Delta = 3$ ,  $H_\Delta \simeq 1.833$ . For  $\Delta > 3$ , CKR-ROUNDING gives a worse than 2 bound while we give an alternate rounding which gives a  $2(1 - 1/k)$ -approximation. Our analysis of CKR-ROUNDING differs from that in [2] since we cannot use the “edge alignment” properties of the fractional solution that hold for graphs and were exploited in [2]; our analysis for HYPERGRAPH-MC is inspired by the proof given by Williamson and Shmoys [25].

It is natural to wonder whether CKR-ROUNDING is crucial to obtaining a bound that is better than 2 for these problems, and in particular whether it gives a 1.5-approximation for SYM-SUB-MP. We show that a  $1.5 - 1/k$ -approximation for HYPERGRAPH-MP (and hence GRAPH-MC also) can be obtained via a different algorithm as well; in particular, the crucial ingredient in CKR-ROUNDING for GRAPH-MC when viewed as a special case of HYPERGRAPH-MP is the correlation provided by the use of the same  $\theta$  for all  $i$ ; one can

replace the random permutation by the uncrossing scheme in Lemma 3.1. We describe this algorithm in the next section. However, for HYPERGRAPH-MC, the random permutation is important in proving the  $H_\Delta$ -bound.

### 3.1 A 1.5-approximation for HYPERGRAPH MULTIWAY PARTITION

We start by understanding the objective function of SUBMP-REL in the context of HYPERGRAPH-MP. Let  $\mathbf{x}$  be a feasible fractional solution, and let  $\mathbf{x}_i = (x(v_1, i), \dots, x(v_n, i))$  be the allocation to  $i$ . Recall that  $f$  here is the hypergraph cut function. What is  $\sum_{i=1}^n \hat{f}(\mathbf{x}_i)$ ? For each terminal  $i$  and each hyperedge  $e$ , let  $I(e, i) = [\min_{v \in e} x(v, i), \max_{v \in e} x(v, i)]$ . Let  $d(e, i)$  denote the length of  $I(e, i)$ , and let  $d(e) = \sum_{i=1}^k d(e, i)$ . Note that  $d(e) \in [0, |e|]$ .

**Lemma 3.3.**  $\sum_{i=1}^k \hat{f}(\mathbf{x}_i) = \sum_e w(e) d(e)$ .

**Proof:** Consider a hyperedge  $e$ . Let  $A(i, \theta)$  be the set whose characteristic vector is  $\mathbf{x}_i^\theta$ . For each  $\theta \in [0, \min_{v \in e} x(v, i)]$ , the set  $A(i, \theta)$  contains all the vertices of  $e$ , and thus  $e \notin \delta(A(i, \theta))$ . For each  $\theta \in (\min_{v \in e} x(v, i), \max_{v \in e} x(v, i)]$ , the set  $A(i, \theta)$  contains at least one vertex of  $e$  but not all of the vertices of  $e$ , and thus  $e \in \delta(A(i, \theta))$ . Finally, for each  $\theta \in (\max_{v \in e} x(v, i), 1]$ , the set  $A(i, \theta)$  does not contain any vertex of  $e$ , and thus  $e \notin \delta(A(i, \theta))$ . Therefore the contribution of  $e$  to  $\hat{f}(\mathbf{x}_i)$  is equal to  $(\max_{v \in e} x(v, i) - \min_{v \in e} x(v, i))w(e) = d(e, i)w(e)$ .  $\square$

A crucial technical lemma that we need is the following which states that the contribution of any  $i$  to  $d(e)$  is at most  $d(e)/2$ .

**Lemma 3.4.** For any  $i$ ,  $d(e, i) \leq d(e)/2$ .

**Proof:** Let  $u = \operatorname{argmax}_{v \in e} x(v, i)$ , and  $w = \operatorname{argmin}_{v \in e} x(v, i)$ . We have

$$\begin{aligned} d(e, i) &= x(u, i) - x(w, i) \\ &= \left(1 - \sum_{j \neq i} x(u, j)\right) - \left(1 - \sum_{j \neq i} x(w, j)\right) \\ &= \sum_{j \neq i} (x(w, j) - x(u, j)) \\ &\leq \sum_{j \neq i} \left(\max_{v \in e} x(v, j) - \min_{v \in e} x(v, j)\right) \\ &= \sum_{j \neq i} d(e, j) \\ &= d(e) - d(e, i) \end{aligned}$$

Therefore  $d(e, i) \leq d(e)/2$ .  $\square$

The algorithm SYMSUBMP-ROUNDING that we analyze is described below. We can prove that CKR-ROUNDING gives the same bound; however, SYMSUBMP-ROUNDING and its analysis are perhaps more intuitive in the context of symmetric functions. The algorithm does  $\theta$ -rounding to obtain sets  $A(1, \theta), \dots, A(k, \theta)$  and then uncrosses these sets to make them disjoint without increasing the expected cost (see Lemma 3.1).

**SYMSUBMP-ROUNDING**

```

let  $x$  be a feasible solution to SUBMP-REL
pick  $\theta \in [0, 1]$  uniformly at random
 $A(i, \theta) \leftarrow \{v \mid x(v, i) \geq \theta\}$  for each  $i$  ( $1 \leq i \leq k$ )
 $\langle\langle \text{uncross } A(1, \theta), \dots, A(k, \theta) \rangle\rangle$ 
 $A'_i \leftarrow A(i, \theta)$  for each  $i$  ( $1 \leq i \leq k$ )
while there exist  $i \neq j$  such that  $A'_i \cap A'_j \neq \emptyset$ 
    if  $(f(A'_i) + f(A'_j - A'_i) \leq f(A'_i) + f(A'_j))$ 
         $A'_j \leftarrow A'_j - A'_i$ 
    else
         $A'_i \leftarrow A'_i - A'_j$ 
return  $(A'_1, \dots, A'_{k-1}, V - (A'_1 \cup \dots \cup A'_{k-1}))$ 

```

**Theorem 3.5.** SYMSUBMP-ROUNDING achieves an 1.5-approximation for HYPERGRAPH-MP.

We remark that we can change the algorithm and the analysis slightly to achieve a  $(1.5 - 1/k)$ -approximation; we give the details in Appendix D.

**Lemma 3.6.** Let  $i^*$  be the index such that the interval  $I(e, i^*)$  has the rightmost ending point among the intervals  $I(e, i)$ . More precisely,  $I(e, i^*)$  is an interval such that  $\max_{v \in e} x(v, i^*) = \max_i \max_{v \in e} x(v, i)$ ; if there are several such intervals, we choose one arbitrarily. Let  $Z_e$  be an indicator random variable equal to 1 iff  $e \in \delta(V - (A(1, \theta) \cup \dots \cup A(k, \theta)))$ . Then  $\mathbb{E}[Z_e] \leq d(e, i^*)$ .

**Proof:** Note that  $Z_e$  is equal to 1 only if (1) for any terminal  $i$ ,  $\theta$  is at least  $\min_{v \in e} x(v, i)$  and (2) there exists a terminal  $\ell$  such that  $\theta \in I(e, \ell)$ . If there exists a terminal  $i$  such that  $\theta$  is smaller than  $\min_{v \in e} x(v, i)$ , all of the vertices of  $e$  are in  $A(i, \theta)$ . If there does not exist a terminal  $\ell$  such that  $\theta \in I(e, \ell)$ , either all of the vertices of  $e$  are in  $A(i, \theta)$  for some  $i$  or all of the vertices of  $e$  are in  $V - (A(1, \theta) \cup \dots \cup A(k, \theta))$ . Finally, we note that (1) and (2) imply that  $\theta$  is in  $I(e, i^*)$ . By (2),  $\theta$  is at most  $\max_{v \in e} x(v, i^*)$  and, by (1),  $\theta$  is at least  $\min_{v \in e} x(v, i^*)$ .  $\square$

**Proof of Theorem 3.5:** It follows from the sub-additivity of  $f$  and Lemma 3.1 that the cost of the partition returned by SYMSUBMP-ROUNDING is at most

$$\sum_{i=1}^k f(A'_i) + f(V - (A'_1 \cup \dots \cup A'_k)) \leq \sum_{i=1}^k f(A(i, \theta)) + f(V - (A(1, \theta) \cup \dots \cup A(k, \theta)))$$

Let  $\text{OPT}_{\text{FRAC}} = \sum_{i=1}^k \hat{f}_s(\mathbf{x}_i)$ . By Lemma 3.6 and Lemma 3.4,

$$\mathbb{E}[f(V - (A(1, \theta) \cup \dots \cup A(k, \theta)))] \leq \sum_e \frac{w(e)d(e)}{2} = \frac{\text{OPT}_{\text{FRAC}}}{2}$$

Finally,  $\mathbb{E}[\sum_{i=1}^k f(A(i, \theta))] = \text{OPT}_{\text{FRAC}}$ , and therefore the expected cost of the allocation is at most  $1.5\text{OPT}_{\text{FRAC}}$ .  $\square$

### 3.2 Algorithms for HYPERGRAPH MULTIWAY CUT

Now we consider HYPERGRAPH-MC. For each hyperedge  $e$ , pick an arbitrary representative node  $r(e) \in e$ . Define the function  $f : 2^V \rightarrow \mathbb{R}_+$  as follows: for  $A \subseteq V$ , let  $f(A) = \sum_{e: r(e) \in A, e \not\subseteq A} w(e)$  be the weight of hyperedges whose representatives are in  $A$  and they cross  $A$ . It is easy to verify that  $f$  is asymmetric and submodular. SUB-MP with this function  $f$  captures HYPERGRAPH-MC [18].

Let  $\mathbf{x}$  be a feasible fractional allocation and  $\mathbf{x}_i$  be the allocation for  $i$ . For each hyperedge  $e$  and each terminal  $i$ , let  $I(e, i) = [\min_{v \in e} x(v, i), \max_{v \in e} x(v, i)]$ . Let  $d(e, i) = x(r(e), i) - \min_{v \in e} x(v, i)$  and  $d(e) = \sum_{i=1}^k d(e, i)$ .

**Lemma 3.7.**  $\sum_{i=1}^k \hat{f}(\mathbf{x}_i) = \sum_e w(e)d(e)$ .

**Proof:** Consider a hyperedge  $e$ . Let  $A(i, \theta)$  be the set whose characteristic vector is  $\mathbf{x}_i^\theta$ . For each  $\theta \in [0, \min_{v \in e} x(v, i)]$ , the set  $A(i, \theta)$  contains all the vertices of  $e$  and therefore  $e \notin \delta(A(i, \theta))$ . For each  $\theta \in (\min_{v \in e} x(v, i), x(r(e), i)]$ , the set  $A(i, \theta)$  contains the representative  $r(e)$  of  $e$  and  $e \in \delta(A(i, \theta))$ . Finally, for each  $\theta \in (x(r(e), i), 1]$ , the set  $A(i, \theta)$  does not contain the representative  $r(e)$ . Therefore the contribution of  $e$  to  $f(\mathbf{x}_i)$  is equal to  $(x(r(e), i) - \min_{v \in e} x(v, i))w(e) = d(e, i)w(e)$ .  $\square$

**HALF-ROUNDING**

let  $x$  be a solution to SUBMP-REL  
 pick  $\theta \in (1/2, 1]$  uniformly at random  
 for  $i = 1$  to  $k - 1$   
      $A(i, \theta) \leftarrow \{v \mid x(v, i) \geq \theta\}$   
      $U(\theta) \leftarrow V - (A(1, \theta) \cup \dots \cup A(k - 1, \theta))$   
 return  $(A(1, \theta), \dots, A(k - 1, \theta), U(\theta))$

The main goal of this section is to show that HALF-ROUNDING is a 2-approximation, and CKR-ROUNDING is a  $H_\Delta$ -approximation for HYPERGRAPH-MC, where  $\Delta$  is the maximum hyperedge size. We remark that we can show that the integrality gap of LE-REL is at most  $2(1 - 1/k)$  using a connection between the distance LP for NODE-WT-MC considered in [9] and LE-REL for HYPERGRAPH-MC; we give the details in Appendix E.

**Lemma 3.8.** *Let  $e$  be any hyperedge, and let  $z$  be any vertex in  $e$ . Let  $R(z) = \{i \mid x(z, i) = \max_{v \in e} x(v, i)\}$ . Then  $\sum_{i \in R(z)} |I(e, i)| \leq d(e)$ .*

**Proof:** Let  $u$  be the representative of  $e$ . If  $z = u$ , the lemma is immediate. Therefore we may assume that  $z \neq u$ . We partition  $\{1, 2, \dots, k\}$  into two sets: the set  $S(z)$  consisting of all coordinates  $i$  such that  $x(z, i)$  is smaller than  $x(u, i)$ , and the set  $B(z)$  consisting of all coordinates  $i$  such that  $x(z, i)$  is at least  $x(u, i)$ . Since  $\sum_{i=1}^k x(z, i)$  and  $\sum_{i=1}^k x(u, i)$  are both equal to 1, the total difference between  $x(u, i)$  and  $x(z, i)$  over all coordinates  $i \in S(z)$  is equal to the total difference between  $x(z, i)$  and  $x(u, i)$  over all coordinates  $i \in B(z)$ . Therefore

$$\begin{aligned} \sum_{i \in B(z)} (x(z, i) - x(u, i)) &= \sum_{i \in S(z)} (x(u, i) - x(z, i)) \\ &\leq \sum_{i \in S(z)} \left( x(u, i) - \min_{v \in e} x(v, i) \right) \\ &\leq d(e) - \sum_{i \in B(z)} \left( x(u, i) - \min_{v \in e} x(v, i) \right) \end{aligned}$$

Since  $R(z)$  is a subset of  $B(z)$ , the lemma follows.  $\square$

**Corollary 3.9.** *For each  $i$ , the length of the interval  $I(e, i)$  is at most  $d(e)$ .*

**Proof:** Let  $\beta = \max_{i=1}^k \max_{v \in e} x(v, i)$ . Let  $s_\ell$  be a terminal and let  $b$  be a vertex in  $e$  such that  $x(b, \ell) = \beta$ . Since  $\ell$  is in  $R(b)$ , the corollary follows from Lemma 3.8.  $\square$

**Theorem 3.10.** *Let  $F$  be the set of all hyperedges crossing the partition returned by HALF-ROUNDING. For each hyperedge  $e$ ,  $\Pr[e \in F] \leq 2d(e)$ .*

**Proof:** Let  $I(e, i^*)$  be the interval with the rightmost right interval among the intervals  $I(e, 1), \dots, I(e, k)$ ; if there are several such intervals, we pick one arbitrarily. Note that  $e$  is in  $F$  only if  $\theta$  is in the interval  $I(e, i^*)$ . Therefore the probability that  $e$  is in  $F$  is at most  $2|I(e, i^*)|$ . By Corollary 3.9, the length of  $I(e, i^*)$  is at most  $d(e)$ .  $\square$

**Theorem 3.11.** *Let  $F$  be the set of all hyperedges crossing the partition returned by CKR-ROUNDING. For each hyperedge  $e$ ,  $\Pr[e \in F] \leq H_{|e|} \cdot d(e)$ .*

**Proof Sketch:** We say that  $s_i$  *splits*  $e$  if  $\theta \in I(e, i)$ . Let  $X_i$  be the event that  $s_i$  splits  $e$ . We say that  $s_i$  *touches*  $e$  if  $\theta \leq \max_{v \in e} x(v, i)$ . (Note that  $\max_{v \in e} x(v, i)$  is the right endpoint of the interval  $I(e, i)$ .) We say that  $s_i$  *settles*  $e$  if  $s_i$  is the first terminal in the permutation  $\pi$  that touches  $e$ . Let  $Y_i$  be the event that  $s_i$  settles  $e$ .

Note that the edge  $e$  is in  $F$  only if there is a terminal  $s_i$  that splits *and* settles  $e$ . Therefore we can upper bound the probability that  $e$  is in  $F$  by  $\sum_{i=1}^k \Pr[X_i \wedge Y_i]$ .

We relabel the terminals so that the ordering of the intervals  $\{I(e, i)\}_{1 \leq i \leq k}$  from right to left according to their ending point is  $I(e, 1), I(e, 2), \dots, I(e, k)$ . (If there are several intervals with the same ending point, we break ties arbitrarily.) After relabeling the intervals,  $s_i$  settles  $e$  only if, for each  $j < i$ ,  $\pi(s_i) < \pi(s_j)$ . This observation, together with the fact that  $s_i$  splits  $e$  with probability  $|I(e, i)|$ , implies that  $\Pr[X_i \wedge Y_i] \leq |I(e, i)|/i$ .

Finally, let  $L(z) = \{i \mid x(z, i) = \max_{v \in e} x(v, i)\}$ . If an index  $i$  belongs to more than one set  $L(z)$ , we only add  $i$  to one of the sets (chosen arbitrarily). Note that, by Lemma 3.8, the total length of the intervals  $I(e, i)$  where  $i \in L(z)$  is at most  $d(e)$ . This, together with the fact that the sets  $L(z)$  are disjoint and their union is  $\{1, 2, \dots, k\}$ , implies that  $\sum_{i=1}^k \Pr[X_i \wedge Y_i]$  is at most  $H_{|e|} \cdot d(e)$ .  $\square$

**Proposition 3.12.** *The analysis in Theorem 3.11 is tight.*

**Proof:** Let  $e$  be a hyperedge with representative  $u$ . Let  $\epsilon \in (0, 1)$  be such that  $\epsilon|e| \leq 1$ . Consider a solution  $x$  that assigns the following values to the vertices of  $e$ . For each terminal  $i > |e|$  and each vertex  $z \in e$ , we have  $x(z, i) = 0$ . For each terminal  $i$  such that  $1 < i \leq |e|$ , we have  $x(u, i) = (|e| - i)\epsilon$ . Finally,  $x(u, 1) = 1 - \sum_{i=2}^k x(u, i)$ . Let  $v_2, \dots, v_{|e|}$  denote the remaining vertices of  $e$  (other than  $u$ ). Now consider an index  $j$  such that  $2 \leq j \leq |e|$ . We have  $x(v_j, 1) = \epsilon$ ,  $x(v_j, j) = x(u, j) - \epsilon$ , and  $x(v_j, i) = x(u, i)$  for all  $i \neq j$ . Note that  $\sum_{i=1}^k x(v, i)$  is equal to 1 for all vertices  $v \in e$  and the distance  $d(e)$  is equal to  $\epsilon$ . It is straightforward to verify that  $e$  is in  $F$  with probability at least  $H_{|e|} \cdot d(e)$ .  $\square$

## 4 Submodular Cost Labeling

In this section we consider SUB-LABEL, which generalizes MONOTONE-MSCA, uniform metric labeling, hub location, and other problems. A natural algorithm here is KT-ROUNDING, which we have already introduced in Section 2. We also describe a different algorithm, SYMSUBLABEL-ROUNDING, which is appropriate for SUB-LABEL when the cut function is an arbitrary symmetric submodular function. We obtain several results that we state below.

The next two results consider the SUB-LABEL problem on hypergraphs in which  $h$  is the following function. For each edge hyperedge  $e$ , pick an arbitrary representative node  $r(e) \in e$ . We define the function  $h : 2^V \rightarrow \mathbb{R}_+$  as follows: for  $A \subseteq V$ , let  $f(A) = \sum_{e: r(e) \in A, e \not\subseteq A} w(e)$  be the weight of hyperedges whose representatives are in  $A$  and they cross  $A$ . We refer to this function as the hypergraph separation cost function.

**Theorem 4.1.** *If  $h$  is the hypergraph separation cost function and each  $g_i$  is modular, KT-ROUNDING achieves a  $\Delta$ -approximation for SUB-LABEL.*

**Theorem 4.2.** *If  $h$  is the hypergraph separation cost function and each  $g_i$  is a monotone submodular function, KT-ROUNDING achieves an  $O(\ln n + \Delta)$  approximation for SUB-LABEL.*

**SYMSUBLABEL-ROUNDING**

```

let  $x$  be a solution to LE-REL
 $S \leftarrow \emptyset$      $\langle\langle$ set of all assigned vertices $\rangle\rangle$ 
 $A_i \leftarrow \emptyset$  for all  $i$  ( $1 \leq i \leq k$ )
while  $S \neq V$ 
    pick  $i \in \{1, 2, \dots, k\}$  uniformly at random
    pick  $\theta \in [0, 1]$  uniformly at random
     $A_i \leftarrow A_i \cup \{v \mid x(v, i) \geq \theta\}$ 
     $S \leftarrow S \cup \{v \mid x(v, i) \geq \theta\}$ 
 $\langle\langle$ uncross  $A_1, \dots, A_k$  $\rangle\rangle$ 
 $A'_i \leftarrow A_i$  for all  $i$  ( $1 \leq i \leq k$ )
while there exist  $i \neq j$  such that  $A'_i \cap A'_j \neq \emptyset$ 
    if  $(f(A'_i) + f(A'_j - A'_i) \leq f(A'_i) + f(A'_j))$ 
         $A'_j \leftarrow A'_j - A'_i$ 
    else
         $A'_i \leftarrow A'_i - A'_j$ 
return  $(A'_1, \dots, A'_k)$ 

```

**Theorem 4.3.** *If  $h$  is a symmetric submodular function and each  $g_i$  is a monotone submodular function, SYMSUBLABEL-ROUNDING achieves an  $O(\ln n)$  approximation for SUB-LABEL.*

Let  $\text{PART}_{\text{FRAC}} = \sum_{i=1}^k \hat{h}(\mathbf{x}_i)$  be the partition cost of LE-REL, and let  $\text{COST}_{\text{FRAC}} = \sum_{i=1}^k \hat{g}_i(\mathbf{x}_i)$  be the assignment cost of LE-REL.

Consider the SUB-LABEL problem in which  $h$  is the hypergraph separation cost function. For each hyperedge  $e$ , let  $d(e) = \sum_{i=1}^k (x(r(e), i) - \min_{v \in e} x(v, i))$ . By Lemma 3.7,  $\text{PART}_{\text{FRAC}} = \sum_e w(e)d(e)$ . In order to bound the expected partition cost of the labeling constructed by KT-ROUNDING, we consider each hyperedge separately, and we give an upper bound on the probability that the hyperedge has at least two vertices with different labels. We say that a hyperedge  $e$  is split in some iteration of KT-ROUNDING if there exists an iteration  $\ell$  such that at least one vertex of  $e$  is assigned a label in iteration  $\ell$  but not all vertices of  $e$  are assigned a label in iteration  $\ell$ . The following lemma gives an upper bound on the probability that a hyperedge  $e$  is split.

**Lemma 4.4.** *For each hyperedge  $e$ , the probability that  $e$  is split is at most  $\Delta d(e)$ .*

Using Lemma 4.4, we can complete the proofs of Theorem 4.1 and Theorem 4.2 as follows.

**Proof of Theorem 4.1:** Let  $(A_1, \dots, A_k)$  be the partition returned by KT-ROUNDING, and let  $\text{PART}_{\text{INT}} = \sum_{i=1}^k h(A_i)$  and  $\text{COST}_{\text{INT}} = \sum_{i=1}^k g_i(A_i)$ . As shown in [15], KT-ROUNDING assigns label  $i$  to  $v$  with probability  $x(v, i)$ . Thus  $\mathbb{E}[\text{COST}_{\text{INT}}] = \text{COST}_{\text{FRAC}}$ .

By Lemma 3.7,  $\text{PART}_{\text{FRAC}} = \sum_e w(e)d(e)$ . Therefore, by Lemma 4.4,  $\mathbb{E}[\text{PART}_{\text{INT}}] \leq \Delta \cdot \text{PART}_{\text{FRAC}}$ .  $\square$

**Proof of Theorem 4.2:** Let  $(A_1, \dots, A_k)$  be the partition returned by KT-ROUNDING, and let  $\text{PART}_{\text{INT}} = \sum_{i=1}^k h(A_i)$  and  $\text{COST}_{\text{INT}} = \sum_{i=1}^k g_i(A_i)$ . Using the argument in the proof of Theorem 2.6, we can show that  $\mathbb{E}[\text{COST}_{\text{INT}}] \leq O(\ln n) \text{COST}_{\text{FRAC}}$ . Additionally, by Lemma 4.4,  $\mathbb{E}[\text{PART}_{\text{INT}}] \leq \Delta \text{PART}_{\text{FRAC}}$ .  $\square$

Now we turn our attention to the proof of Lemma 4.4.

**Proof of Lemma 4.4:** Consider iteration  $\ell$  of KT-ROUNDING, and let  $i_\ell$  and  $\theta_\ell$  be the label and threshold in iteration  $\ell$ . We say that iteration  $\ell$  *cuts*  $e$  if  $\theta_\ell \in [\min_{v \in e} x(v, i_\ell), \max_{v \in e} x(v, i_\ell)]$ . We say that iteration  $\ell$  *touches*  $e$  if  $\theta_\ell$  is in the interval  $[0, \max_{v \in e} x(v, i_\ell)]$ . Let  $X_\ell$  and  $Z_\ell$  be the events that  $\ell$  cuts and touches  $e$  (respectively). The probability that  $e$  is split in iteration  $\ell$  is at most  $\Pr[X_\ell] / \Pr[Z_\ell]$ . We have

$$\Pr[X_\ell] \leq \frac{1}{k} \sum_{i=1}^k \left( \max_{v \in e} x(v, i) - \min_{v \in e} x(v, i) \right) \leq \frac{\Delta d(e)}{k}$$

where the last inequality follows from Lemma 3.8. Additionally,

$$\Pr[Z_\ell] = \frac{1}{k} \sum_{i=1}^k \max_{v \in e} x(v, i) \geq \frac{1}{k}$$

The last inequality follows from the fact that, for any vertex  $w \in e$ ,  $\sum_{i=1}^k x(w, i) = 1$ . It follows that the probability that  $e$  is split in iteration  $j$  is at most  $\Delta d(e)$ .  $\square$

**Proof of Theorem 4.3:** Let  $i_\ell$  and  $\theta_\ell$  be the label and  $\theta$  value chosen in the  $\ell$ -th iteration of the *first* while loop SYMSUBLABEL-ROUNDING. For each  $i$ , let  $\mathcal{A}_i = \cup_{\ell: i_\ell=i} \{v \mid x(v, i) \geq \theta_\ell\}$ . Let  $\text{PART}_{\text{BALLS}} = \sum_{i=1}^k f(\mathcal{A}_i)$  and  $\text{COST}_{\text{BALLS}} = \sum_{i=1}^k g_i(\mathcal{A}_i)$ . (Note that  $\mathcal{A}_i$  is the set  $A_i$  at the end of the *first* while loop of SYMSUBLABEL-ROUNDING.)

Using the argument in the proof of Theorem 2.6, we can show that

$$\mathbb{E}[\text{COST}_{\text{BALLS}}] \leq O(\ln n) \text{COST}_{\text{FRAC}}$$

and

$$\mathbb{E}[\text{PART}_{\text{BALLS}}] \leq O(\ln n) \text{PART}_{\text{FRAC}}$$

Let  $(A'_1, \dots, A'_k)$  be the partition returned by SYMSUBLABEL-ROUNDING. Let  $\text{PART}_{\text{INT}} = \sum_{i=1}^k f(A'_i)$  and  $\text{COST}_{\text{INT}} = \sum_{i=1}^k g_i(A'_i)$ . By Lemma 3.1,

$$\mathbb{E}[\text{PART}_{\text{INT}}] \leq \mathbb{E}[\text{PART}_{\text{BALLS}}] \leq O(\ln n) \text{PART}_{\text{FRAC}}$$

Since each  $g_i$  is monotone,

$$\mathbb{E}[\text{COST}_{\text{INT}}] \leq \mathbb{E}[\text{COST}_{\text{BALLS}}] \leq O(\ln n) \text{COST}_{\text{FRAC}}.$$

$\square$

**Integrality Gap Example:** We remark that we can generalize the integrality gap example of [15] in order to show that the integrality gap of LE-REL is at least  $\Delta(1 - 1/k)$  for SUB-LABEL when  $h$  is the hypergraph separation cost function, even if each  $g_i$  is modular.

Consider a  $\Delta$ -uniform complete hypergraph on  $k$  vertices; all  $\binom{k}{\Delta}$  edges are present, and each edge has unit weight. For each vertex  $i$ , the cost of assigning label  $j$  to  $i$  is zero if  $i \neq j$ , and infinity otherwise.

It is easy to see that an optimal integral solution picks a label  $i$  and assigns label  $i$  to all vertices except  $i$ , and it assigns some other label to  $i$ . Thus the integral optimum is  $\binom{k-1}{\Delta-1}$ . Setting  $x(i, j) = 1/(k-1)$  for all  $i \neq j$  gives us a fractional solution of cost  $\binom{k}{\Delta}/(k-1)$ .

**Acknowledgments:** We thank Lisa Fleischer for suggesting that we contact Zoya Svitkina about MSCA and thank Zoya for pointing out her work in [22] on monotone MSCA. CC thanks Jan Vondrak for pointing out the interpretation of the Lovász extension from his paper [24] which was very helpful in thinking about rounding procedures. AE thanks Sungjin Im and Ben Moseley for discussions.

## References

- [1] Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a submodular set function subject to a matroid constraint (extended abstract). In *Integer Programming and Combinatorial Optimization (IPCO)*, pages 182–196, 2007.
- [2] Gruia Calinescu, Howard J. Karloff, and Yuval Rabani. An improved approximation algorithm for multiway cut. *Journal of Computer and System Sciences*, 60(3):564–574, 2000. Preliminary version in STOC 1998.
- [3] Chandra Chekuri and Alina Ene. Approximation algorithms for submodular multiway partition. Manuscript, April 2011.
- [4] Elias Dahlhaus, David S. Johnson, Christos H. Papadimitriou, Paul D. Seymour, and Mihalis Yannakakis. The complexity of multiterminal cuts. *SIAM Journal on Computing*, 23(4):864–894, 1994. Preliminary version in STOC 1992.
- [5] Andrew DeLong, Anton Osokin, Hossam N. Isack, and Yuri Boykov. Fast approximate energy minimization with label costs. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, pages 2173–2180, 2010.
- [6] Shaddin Dughmi. Submodular functions: Extensions, distributions, and algorithms. A survey. *CoRR*, abs/0912.0322, 2009.
- [7] Ari Freund and Howard J. Karloff. A lower bound of  $8/(7+(1/k)-1)$  on the integrality ratio of the calinescu-karloff-rabani relaxation for multiway cut. *Information Processing Letters*, 75(1-2):43–50, 2000.
- [8] Takuro Fukunaga. Computing Minimum Multiway Cuts in Hypergraphs from Hypertree Packings. *Integer Programming and Combinatorial Optimization (IPCO)*, pages 15–28, 2010.
- [9] Naveen Garg, Vijay V. Vazirani, and Mihalis Yannakakis. Multiway cuts in node weighted graphs. *Journal of Algorithms*, 50(1):49–61, 2004. Preliminary version in ICALP 1994.
- [10] Dongdong Ge, Yinyu Ye, and Jiawei Zhang. The Fixed-Hub Single Allocation Problem: A Geometric Rounding Approach. Preprint available at <http://www.stanford.edu/~yyye/revisedHub.pdf>, 2007.
- [11] Gagan Goel, Chinmay Karande, Pushkar Tripathi, and Lei Wang. Approximability of combinatorial problems with multi-agent submodular cost functions. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 755–764, 2009.
- [12] Michel X. Goemans, Nicholas J. A. Harvey, Satoru Iwata, and Vahab S. Mirrokni. Approximating submodular functions everywhere. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 535–544, 2009.
- [13] Satoru Iwata and Kiyohito Nagano. Submodular function minimization under covering constraints. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 671–680, 2009.
- [14] David R. Karger, Philip N. Klein, Clifford Stein, Mikkell Thorup, and Neal E. Young. Rounding algorithms for a geometric embedding of minimum multiway cut. *Mathematics of Operations Research*, 29(3):436–461, 2004. Preliminary version in STOC 1999.
- [15] Jon M. Kleinberg and Éva Tardos. Approximation algorithms for classification problems with pairwise relationships: Metric labeling and Markov random fields. *Journal of the ACM (JACM)*, 49(5):616–639, 2002. Preliminary version in FOCS 1999.



- [16] Eugene L. Lawler. Cutsets and partitions of hypergraphs. *Networks*, 3(3):275–285, 1973.
- [17] László Lovász. Submodular functions and convexity. *Mathematical programming: the state of the art*, pages 235–257, 1983.
- [18] Kazumasa Okumoto, Takuro Fukunaga, and Hiroshi Nagamochi. Divide-and-conquer algorithms for partitioning hypergraphs and submodular systems. *Algorithmica*, pages 1–20, 2010. Preliminary version in ISAAC 2009.
- [19] Maurice Queyranne. Minimizing symmetric submodular functions. *Mathematical Programming*, 82(1):3–12, 1998. Preliminary version in SODA 1995.
- [20] Zoya Svitkina and Lisa Fleischer. Submodular approximation: Sampling-based algorithms and lower bounds. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 697–706, 2008.
- [21] Zoya Svitkina and Éva Tardos. Min-max multiway cut. *Approximation, Randomization, and Combinatorial Optimization (APPROX-RANDOM)*, pages 207–218, 2004.
- [22] Zoya Svitkina and Éva Tardos. Facility location with hierarchical facility costs. *ACM Transactions on Algorithms (TALG)*, 6(2):1–22, 2010. Preliminary version in SODA 2006.
- [23] Jan Vondrák. Optimal approximation for the submodular welfare problem in the value oracle model. In *ACM Symposium on Theory of Computing (STOC)*, pages 67–74, 2008.
- [24] Jan Vondrák. Symmetry and Approximability of Submodular Maximization Problems. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 651–670, 2010.
- [25] David P. Williamson and David B. Shmoys. The design of approximation algorithms. Preprint available at <http://www.designofapproxalgs.com>, 2010.
- [26] Mingyu Xiao. Finding minimum 3-way cuts in hypergraphs. *Information Processing Letters*, 110(14-15):554–558, 2010. Preliminary version in TAMC 2008.
- [27] Liang Zhao, Hiroshi Nagamochi, and Toshihide Ibaraki. Greedy splitting algorithms for approximating multiway partition problems. *Mathematical Programming*, 102(1):167–183, 2005.

## A Definition of the Lovász extension

Let  $f : \{0, 1\}^n \rightarrow \mathbb{R}$  be a function. The Lovász extension  $\hat{f}$  of  $f$  is the function  $\hat{f} : [0, 1]^n \rightarrow \mathbb{R}$  defined as follows. Let  $\mathbf{x}$  be a vector in  $[0, 1]^n$ . We relabel the vertices as  $1, 2, \dots, n$  so that  $x_1 \geq x_2 \geq \dots \geq x_n$ ; for ease of notation, let  $x_0 = 1$  and  $x_{n+1} = 0$ . Let  $S_i = \{1, 2, \dots, i\}$ . The value of  $\hat{f}$  at  $\mathbf{x}$  is equal to

$$\hat{f}(\mathbf{x}) = \sum_{i=0}^n (x_{i+1} - x_i) f(S_i)$$

It is straightforward to verify that  $\sum_{i=0}^n (x_{i+1} - x_i) f(S_i) = \mathbb{E}_{\theta \in [0, 1]} [f(\mathbf{x}^\theta)]$ .

Another useful extension for a function  $f$  is its convex closure, which is defined as follows. For each set  $S \subseteq V$ , we let  $\mathbf{1}_S$  denote the characteristic vector of  $S$ ; that is, the  $i$ -th coordinate of  $\mathbf{1}_S$  is equal to 1 if  $i$  is in  $S$  and 0 otherwise. The convex closure  $f^-$  of  $f$  is the function  $f^- : [0, 1]^n \rightarrow \mathbb{R}$  where  $f^-(\mathbf{x}) = \min \{ \sum_{S \subseteq V} \lambda_S f(S) : \sum_{S \subseteq V} \lambda_S \mathbf{1}_S = \mathbf{x}, \sum_{S \subseteq V} \lambda_S = 1, \lambda_S \geq 0 \}$ . The Lovász extension  $\hat{f}$  of  $f$  is equal to the convex closure  $f^-$  of  $f$  iff  $f$  is submodular; see for instance [6]. Using this result, we can show that LE-REL can be solved in time that is polynomial in  $n$  and  $\log(\max_{i, S \subseteq V} f_i(S))$  via the ellipsoid method.

Since  $\hat{f}$  is equal to  $f^-$ , we can write LE-REL as follows.

<b>LE-REL-Primal</b>		
min	$\sum_{i=1}^k \sum_{S \subseteq V} \lambda(S, i) f_i(S)$	
	$\sum_{S: v \in S} \lambda(S, i) = x(v, i) \quad \forall v, i$	
	$\sum_{S \subseteq V} \lambda(S, i) = 1 \quad \forall i$	
	$\sum_{i=1}^k x(v, i) = 1 \quad \forall v$	
	$\lambda(S, i) \geq 0 \quad \forall S, i$	
	$x(v, i) \geq 0 \quad \forall v, i$	

**LE-REL-Primal** is an LP with exponentially many variables and polynomially many constraints. Its dual **LE-REL-Dual** has polynomially many variables and exponentially many constraints.

<b>LE-REL-Dual</b>		
max	$\sum_{i=1}^k \beta_i + \sum_{v \in V} \gamma_v$	
	$\sum_{v \in S} \alpha(v, i) + \beta_i \leq f_i(S) \quad \forall S, i$	
	$\gamma_v \leq \alpha(v, i) \quad \forall v, i$	

**Separation oracle for LE-REL-Dual.** Fix an assignment of values to the variables  $\alpha, \beta, \gamma$  in **LE-REL-Dual**. It is easy to check in polynomial time whether  $\gamma_v \leq \alpha(v, i)$  for all  $v, i$  since there are only  $nk$  such constraints. Let  $g_i(S) = \sum_{v \in S} \alpha(v, i) + \beta_i$ . Note that  $g_i$  is a *modular* function and therefore  $f_i - g_i$  is a submodular

function. Using a polynomial time algorithm for submodular function minimization, for a given  $i$ , we can check whether  $f_i(S) - g_i(S) \geq 0$  for all sets  $S \subseteq V$ .

Therefore we can solve **LE-REL-Dual** in time that is polynomial in  $n$  and  $\log(\max_{i,S \subseteq V} f_i(S))$  using the ellipsoid method. Using standard techniques, we can also construct an optimal solution for the primal; we omit the details here.

## B Omitted proofs from Section 2

**Proof of Proposition 2.4:** Consider a terminal  $i$ . For any  $\theta$ ,  $A(i, \theta)$  is a subset of  $\{v \mid v \in V, x(v, i) \geq \theta\}$ . Since  $f_i$  is monotone,  $f_i(A(i, \theta)) \leq f_i(\{v \mid v \in V, x(v, i) \geq \theta\})$ . Therefore

$$\begin{aligned} \mathbb{E}_{\theta \in [0,1]}[A(i, \theta)] &= \int_0^1 f_i(A(i, \theta)) d\theta \\ &\leq \int_0^1 f_i(\{v \mid v \in V, x(v, i) \geq \theta\}) d\theta \\ &= \hat{f}_i(\mathbf{x}_i) \end{aligned}$$

Finally,

$$\mathbb{E}_{i, \theta}[f_i(A(i, \theta))] = \frac{1}{k} \sum_{i=1}^k \mathbb{E}_{\theta \in [0,1]}[f_i(A(i, \theta))] \leq \frac{1}{k} \text{OPT}_{\text{FRAC}}$$

□

**Proof of Proposition 2.5:** Note that

$$\mathbb{E}_{i, \theta}[|A(i, \theta)|] = \frac{1}{k} \sum_{i=1}^k \mathbb{E}_{\theta \in [0,1]}[|A(i, \theta)|]$$

We can prove by induction on the size of  $U$  that  $\sum_{i=1}^k \mathbb{E}_{\theta \in [0,1]}[|A(i, \theta)|]$  is equal to  $|U|$ .

If  $U$  is empty, the claim trivially holds. Therefore we may assume that  $U$  contains at least one element  $z$ . Let  $U' = U - \{z\}$ , and let  $\tilde{x}'$  be the restriction of  $\tilde{x}$  to  $U'$ ; more precisely,  $\tilde{x}'(v, i) = \tilde{x}(v, i)$  for all  $v \in U'$  and all terminals  $i$ . Let  $A'(i, \theta) = \{v \mid v \in U', \tilde{x}'(v, i) \geq \theta\}$ . Note that  $A'(i, \theta)$  is equal to  $A(i, \theta) - \{z\}$  if  $\theta$  is smaller than  $x(z, i)$ , and  $A'(i, \theta)$  is equal to  $A(i, \theta)$  otherwise. Therefore

$$\begin{aligned} \sum_{i=1}^k \mathbb{E}_{\theta \in [0,1]}[|A(i, \theta)|] &= \sum_{i=1}^k \int_0^1 |A(i, \theta)| d\theta \\ &= \sum_{i=1}^k \int_0^{x(z, i)} |A'(i, \theta) \cup \{z\}| d\theta + \sum_{i=1}^k \int_{x(z, i)}^1 |A'(i, \theta)| d\theta \\ &= \sum_{i=1}^k \int_0^1 |A'(i, \theta)| d\theta + \sum_{i=1}^k x(z, i) \\ &= \sum_{i=1}^k \int_0^1 |A'(i, \theta)| d\theta + 1 \\ &= |U'| + 1 \quad (\text{By induction}) \\ &= |U| \end{aligned}$$

□

## C Omitted proofs from Section 3

**Theorem C.1.** *The integrality gap of LE-REL for SYM-SUB-MP is at most  $2(1 - 1/k)$ .*

**Proof:** Let  $\mathbf{x}$  be an optimal solution to LE-REL for a given instance of SYM-SUB-MP. Without loss of generality,  $\hat{f}(\mathbf{x}_k) = \max_i \hat{f}(\mathbf{x}_i)$ . Let  $A(1, \theta), \dots, A(k-1, \theta)$  be sets obtained by applying  $\theta$ -rounding with respect to the first  $k-1$  terminals. By the property of  $\theta$ -rounding, we observe that  $\sum_{i=1}^{k-1} \mathbb{E}[f(A(i, \theta))] = \sum_{i=1}^{k-1} \hat{f}(\mathbf{x}_i) \leq (1 - 1/k) \text{OPT}_{\text{FRAC}}$ . The last inequality follows from the fact that  $\hat{f}(\mathbf{x}_k) \geq \text{OPT}_{\text{FRAC}}/k$ . Note that  $s_i$  belongs only to  $A(i, \theta)$ . We now apply Lemma 3.1 to  $A(1, \theta), \dots, A(k-1, \theta)$  to obtain  $A'_1, \dots, A'_{k-1}$ . We have  $\sum_{i=1}^{k-1} f(A'_i) \leq \sum_{i=1}^{k-1} f(A(i, \theta))$ . Let  $A'_k = V - \cup_{1 \leq i \leq k-1} A'_i$ . By symmetry of  $f$ ,  $f(A'_k) = f(\cup_{1 \leq i \leq k-1} A'_i)$  and, since  $f$  is sub-additive,  $f(A'_k) = f(\cup_{1 \leq i \leq k-1} A'_i) \leq \sum_{i=1}^{k-1} f(A'_i) \leq \sum_{i=1}^{k-1} f(A(i, \theta))$ . We allocate  $A'_k$  to index  $k$  and the total cost of the allocation is upper bounded by  $\sum_i f(A'_i) \leq 2 \sum_{i=1}^{k-1} f(A(i, \theta))$ . Thus the expected cost of the allocation is at most  $2(1 - 1/k) \text{OPT}_{\text{FRAC}}$ . The allocation is feasible since, for each  $i \neq k$ ,  $s_i$  belongs only to  $A(i, \theta)$  and hence to  $A'_i$ , and  $s_k$  belongs to  $A'_k$ .  $\square$

## D Omitted proofs from Subsection 3.1

### SYMSUBMP-ROUNDING

```

let  $x$  be a feasible solution to SUBMP-REL
relabel the terminals so that  $\hat{f}(\mathbf{x}_k) = \max_i \hat{f}(\mathbf{x}_i)$ 
pick  $\theta \in [0, 1]$  uniformly at random
 $A(i, \theta) \leftarrow \{v \mid x(v, i) \geq \theta\}$  for each  $i$  ( $1 \leq i \leq k-1$ )
 $\langle\langle \text{uncross } A(1, \theta), \dots, A(k-1, \theta) \rangle\rangle$ 
 $A'_i \leftarrow A(i, \theta)$  for each  $i$  ( $1 \leq i \leq k-1$ )
while there exist  $i \neq j$  such that  $A'_i \cap A'_j \neq \emptyset$ 
    if  $(f(A'_i) + f(A'_j - A'_i) \leq f(A'_i) + f(A'_j))$ 
         $A'_j \leftarrow A'_j - A'_i$ 
    else
         $A'_i \leftarrow A'_i - A'_j$ 
return  $(A'_1, \dots, A'_{k-1}, V - (A'_1 \cup \dots \cup A'_{k-1}))$ 

```

**Theorem D.1.** SYMSUBMP-ROUNDING achieves an  $(1.5 - 1/k)$ -approximation for HYPERGRAPH-MP.

**Lemma D.2.** *Let  $i^*$  be the index such that the interval  $I(e, i^*)$  has the rightmost ending point among the intervals  $I(e, i)$ , where  $1 \leq i \leq k-1$ . More precisely,  $I(e, i^*)$  is an interval such that  $\max_{v \in e} x(v, i^*) = \max_{1 \leq i \leq k-1} \max_{v \in e} x(v, i)$ ; if there are several such intervals, we choose one arbitrarily. Let  $Z_e$  be an indicator random variable equal to 1 iff  $e \in \delta(V - (A(1, \theta) \cup \dots \cup A(k-1, \theta)))$ . Then  $\mathbb{E}[Z_e] \leq d(e, i^*)$ .*

**Proof:** Note that  $Z_e$  is equal to 1 only if (1) for any terminal  $i \neq k$ ,  $\theta$  is at least  $\min_{v \in e} x(v, i)$  and (2) there exists a terminal  $\ell \neq k$  such that  $\theta \in I(e, \ell)$ . If there exists a terminal  $i \neq k$  such that  $\theta$  is smaller than  $\min_{v \in e} x(v, i)$ , all of the vertices of  $e$  are in  $A(i, \theta)$ . If there does not exist a terminal  $\ell \neq k$  such that  $\theta \in I(e, \ell)$ , either all of the vertices of  $e$  are in  $A(i, \theta)$  for some  $i \neq k$  or all of the vertices of  $e$  are in  $V - (A(1, \theta) \cup \dots \cup A(k-1, \theta))$ . Finally, we note that (1) and (2) imply that  $\theta$  is in  $I(e, i^*)$ : by (2),  $\theta$  is at most  $\max_{v \in e} x(v, i^*)$  and, by (1),  $\theta$  is at least  $\min_{v \in e} x(v, i^*)$ .  $\square$

**Proof of Theorem D.1:** It follows from Lemma 3.1 that the cost of the partition returned by SYMSUBMP-ROUNDING is at most

$$\sum_{i=1}^{k-1} f(A'_i) + f(V - (A'_1 \cup \dots \cup A'_{k-1})) \leq \sum_{i=1}^{k-1} f(A(i, \theta)) + f(V - (A(1, \theta) \cup \dots \cup A(k-1, \theta)))$$

By Lemma D.2 and Lemma 3.4,

$$\mathbb{E}[f(V - (A(1, \theta) \cup \dots \cup A(k-1, \theta)))] \leq \sum_e \frac{w(e)d(e)}{2} = \frac{\text{OPT}_{\text{FRAC}}}{2}$$

Since  $\hat{f}(\mathbf{x}_k) \geq \text{OPT}_{\text{FRAC}}/k$ , we have

$$\mathbb{E} \left[ \sum_{i=1}^{k-1} f(A(i, \theta)) \right] \leq \left(1 - \frac{1}{k}\right) \text{OPT}_{\text{FRAC}}$$

Therefore the expected cost of the allocation is at most  $(1.5 - 1/k)\text{OPT}_{\text{FRAC}}$ .  $\square$

## E Improved integrality gap bound for HYPERGRAPH-MC

The main goal of this section is to establish a connection between the distance LP for NODE-WT-MC considered in [9] and LE-REL for HYPERGRAPH-MC. This connection gives us the following theorem.

**Theorem E.1.** *The integrality gap of HYPERGRAPH-MC is at most  $2(1 - 1/k)$  for instances of HYPERGRAPH-MC.*

The reader may wonder whether we can prove the above theorem directly without recourse to the result from [9]; we believe that it can be done but there are some technical hurdles that we plan to address in a future version of the paper. As we already noted, HYPERGRAPH-MC and NODE-WT-MC are approximation equivalent. Okumoto *et al.* [18] gave an approximation-preserving reduction from HYPERGRAPH-MC to NODE-WT-MC. The reduction maps an instance of HYPERGRAPH-MC to an instance of NODE-WT-MC as follows. Let  $G = (V, E)$  be an instance of HYPERGRAPH-MC, namely a hypergraph with weights on the edges and  $k$  distinguished vertices which we call terminals. We construct a bipartite graph  $H$  as follows. We add all of the vertices of  $G$  to  $H$ , and we give them infinite weight. For each hyperedge  $e$  of  $G$ , we add a vertex to  $H$  of weight  $w_e$ , and we connect it to all of vertices of  $V$  that are contained in the hyperedge. The terminals of  $H$  are the terminals of  $G$ , and it is straightforward to verify that a multiway cut in  $G$  corresponds to a node multiway cut in  $H$  of the same weight and vice-versa.

We remark that there is an approximation-preserving reduction from NODE-WT-MC to HYPERGRAPH-MC as well. Let  $G = (V, E)$  be an instance of NODE-WT-MC, namely a graph with weights on the vertices and  $k$  distinguished vertices called terminals. We may assume without loss of generality that the terminals form an independent set of  $G$ . We subdivide each edge of  $G$ , except the edges incident to the terminals. Now we can view the resulting graph  $G'$  as a bipartite graph with the terminals and the subdividing vertices on the left, and all other vertices on the right. We construct a hypergraph  $H$  as follows: the vertices of  $H$  are the left vertices of  $G'$  and, for each vertex  $v$  on the right,  $H$  has a hyperedge consisting of all neighbors of  $v$  (in the subdivided graph  $G'$ ). The weight of the hyperedge corresponding to  $v$  is  $w(v)$ . The terminals of  $H$  are the terminals of  $G$ , and it is straightforward to verify that a node multiway cut in  $G$  corresponds to a multiway cut in  $H$  of the same weight and vice-versa.

Consider an instance  $G$  of HYPERGRAPH-MC, and let  $\mathbf{x}$  be a feasible solution to LE-REL for this instance. Using the first reduction, we map an instance  $G$  of HYPERGRAPH-MC to the instance  $H$  of NODE-WT-MC described above. In the following, we show that we can map the solution  $\mathbf{x}$  to a solution  $\mathbf{d}$  to the distance LP relaxation for  $H$ . The distance LP has a variable  $d_v$  for each non-terminal  $v$  with the interpretation that  $d_v$  is 1 if  $v$  is in the node multiway cut. Let  $\mathcal{P}_{s_i, s_j}$  denote the set of all paths of  $H$  from  $s_i$  to  $s_j$ .

$$\begin{array}{ll}
\text{(DISTANCE-LP)} & \min \sum_v d_v w_v \\
& \sum_{v \in p} d_v \geq 1 \quad \forall i \neq j, \forall p \in \mathcal{P}_{s_i, s_j} \\
& d_v \geq 0 \quad \forall v \in V(H) - \{s_1, \dots, s_k\}
\end{array}$$

We map the solution  $\mathbf{x}$  to LE-REL to a solution  $\mathbf{d}$  to DISTANCE-LP as follows. For each vertex  $v \in V$ , we set  $d_v = 0$  (recall that all the vertices in  $V$  have infinite weight). Let  $z$  be a vertex of  $H$  that corresponds to the hyperedge  $e$  of  $G$ . Let  $u$  be the representative of  $e$ , and let

$$d_z = \sum_{i=1}^k \left( x(u, i) - \min_{v \in e} x(v, i) \right) = 1 - \sum_{i=1}^k \min_{v \in e} x(v, i)$$

Now consider a path  $p$  of  $H$  between two terminals  $s_a$  and  $s_b$ . Let  $p = v_0 - z_1 - v_1 - z_2 - \dots - z_\ell - v_\ell$ , where  $z_j$  is in  $V(H) - V$ ,  $v_j$  is in  $V$ ,  $v_0 = s_a$ , and  $v_\ell = s_b$ . Let  $e_j$  be the hyperedge corresponding to  $z_j$ ; node that  $e_j$  contains  $v_{j-1}$  and  $v_j$ . For each vertex  $v \in V$ , let  $\bar{x}(v) = (x(v, 1), \dots, x(v, k))$  denote the point on the  $k$ -dimensional simplex to which  $v$  is mapped by the solution  $\mathbf{x}$ , and let  $\|\cdot\|_1$  denote the  $\ell_1$  norm of a vector.

$$\begin{aligned}
\sum_{j=1}^{\ell} d_{z_j} &= \sum_{j=1}^{\ell} \left( 1 - \sum_{i=1}^k \min_{v \in e_j} x(v, i) \right) \\
&\geq \sum_{j=1}^{\ell} \left( 1 - \sum_{i=1}^k \min\{x(v_{j-1}, i), x(v_j, i)\} \right) \quad (e_j \text{ contains } v_{j-1} \text{ and } v_j) \\
&= \sum_{j=1}^{\ell} \sum_{i=1}^k (x(v_{j-1}, i) - \min\{x(v_{j-1}, i), x(v_j, i)\}) \\
&= \sum_{j=1}^{\ell} \sum_{i=1}^k \max\{0, x(v_{j-1}, i) - x(v_j, i)\} \\
&= \frac{1}{2} \sum_{j=1}^{\ell} \sum_{i=1}^k |x(v_{j-1}, i) - x(v_j, i)| \\
&= \frac{1}{2} \|\bar{x}(v_{j-1}) - \bar{x}(v_j)\|_1 \\
&\geq \frac{1}{2} \|\bar{x}(s_a) - \bar{x}(s_b)\|_1 = 1
\end{aligned}$$

Therefore  $\mathbf{d}$  is a feasible solution to DISTANCE-LP. Garg, Vazirany, and Yannakakis [9] showed that the integrality gap of DISTANCE-LP is at most  $2(1 - 1/k)$ . Therefore the integrality gap of LE-REL is at most  $2(1 - 1/k)$  as well. The above argument also establishes that LE-REL is at least as strong a relaxation as DISTANCE-LP for NODE-WT-MC. Easy examples show that it is strictly stronger.